DESCRIPTION

BURST MEMORY ACCESS METHOD TO RECTANGULAR AREA

Technical Field

[0001]   The present invention relates to an information processing device and a data access method which read out a rectangular area inside of image data from a DRAM, the DRAM having a burst-mode which burst-transfers data of successive column addresses; the data access method processes the rectangular area of image data.

Background Art

[0002]   In recent years, multiple functions are being implemented in system LSIs; from the standpoint of system cost reduction, architecture such as unified memory, such as DRAM, are often used. It has also become common for diverse DRAM access requests to be made to a single DRAM.

[0003]   In addition, since system LSIs implemented with multiple functions, such as media processing, demand a high bandwidth, speed enhancement in DRAM is becoming increasingly necessary. Accordingly, memory manufacturers are competing fiercely in DRAM speed enhancement.

[0004]   However, since the operation frequency of a DRAM memory cell itself has not changed from before, the minimum access size to DRAM is rapidly increasing from the perspective of system LSI. As a result, there are no problems transferring long data that is equivalent to a burst length, however there is the problem that when transmitting short data, the transference load of unnecessary data increases and the effective bandwidth will decrease. For example, there is the problem that, for media processing, the effective bandwidth for a motion compensation

process necessary for video decoding will drop; this problem has conventionally been avoided by utilizing high cost DRAM, which tolerates this drop in effective bandwidth.

[0005]   For increasing the effective bandwidth, the image processing device disclosed in Patent Document 1, for example, aims to improve effective bandwidth by including a cache memory for caching single slice data from a frame memory for storing frame data.

[Patent Document 1] Japanese Laid-Open Patent Application No. 2000-175201 Publication

Disclosure of Invention
Problems that Invention is to Solve

[0006]   Since in the past, access to reference images in motion compensation processing has been access on a line basis, in contrast to a rectangular area basis i.e. short data-length lateral access, there is the problem that the transfer load of unnecessary data increases and the effective bandwidth drops.   In other words, since the minimal access size for DRAM is rapidly increasing, as mentioned above, when data shorter than the burst length on the line direction is transferred, there is the problem that the transfer load of unnecessary data increases, and that the effective bandwidth drops.

[0007]   In the conventional address mapping between a physical address of the DRAM and a logical address of image data, in order to earn a burst length for the line access direction, mapping is performed so that the column addresses are successive with the line direction.   As a result, in access to data that is shorter than the burst length on the line direction, such as motion compensation processing, there is the problem that the transfer load of unnecessary data increases and the effective bandwidth drops.

[0008]   Further, even when using a cache memory which caches data on a slice basis, there is the problem that the effective

bandwidth cannot be raised for access to rectangular areas which exceed a boundary of a slice, as in motion compensation processing.

[0009]   The present invention is conceived to solve the aforementioned problems and has as an object to provide an information processing device and data access method which reduce the transfer load of unnecessary data and improve effective bandwidth, even when data is accessed on a rectangular area basis, as in motion compensation processing.

Means to Solve the Problems

[0010]   In order to achieve the above object, the information processing device in the present invention includes a DRAM having a burst mode which burst-transfers data at successive column addresses; one or more data processing units which issue an access request; and an address conversion unit which converts access addresses which are included in the access request issued from the one or more data processing units.   At least one of the one or more data processing units accesses an M x N rectangular area, where M and N are integers, and the address conversion unit converts access addresses so that a column address of data at the $(K+m)$th column, where K and m are integers and $m \leq M$, of an Lth line, and a column address of data at a Kth column of an $(L+n)$th line, where L and n are integers and $n \leq N$, become successive.

[0011]   According to this configuration, there is the effect that the invalid transfer load can be reduced and the effective bandwidth can be improved in a short burst-length access method like rectangular access.

[0012]   Here, some or all areas of the DRAM is a frame memory which stores image data, the rectangular area is M pixels x N lines in the image data, where M and N are integers, and the data processing unit may have a structure that performs one of motion compensation and motion estimation.

[0013]   Here, n may be equal to 2n' (n=2n').

According to this configuration, there is the effect that unnecessary transfer loads can be reduced and the effective bandwidth can be improved in field estimation access.

[0014]   Here, another one of the data processing units may access the image data on a line basis, and may successively read out all data of 2n lines.

[0015]   According to this configuration, there is the effect that invalid transfer loads can be reduced and the effective bandwidth can be improved in DRAM which is mapped more than the address conversion unit above.

[0016]   Here, the data processing unit may decode an inputted stream on a basis of two or more macroblocks, by motion compensation, the DRAM may store the image data decoded by the data processing unit, the information processing device may further include: a memory featuring a smaller storage capacity and faster access speed than the DRAM; a data transfer unit which transfers data from the DRAM to the memory; and the data processing unit accesses the image data stored in the DRAM as reference data.

[0017]   Here, the image data stored in the DRAM may be split into split regions larger in size than the rectangular area, and the data transfer unit may transfer data on a split region basis from the DRAM to the memory, based on the access request from the data processing unit.

[0018]   Here, the data transfer unit may have a structure that includes a register which holds a size of the split region.

[0019]   Since plural rectangular areas can be transferred simultaneously, there is the effect that invalid transfer loads can be reduced and the effective bandwidth can be improved.

[0020]   The data transfer unit may transfer data from the DRAM to the memory when a predetermined number n of read-out requests are outputted from the data processing unit.

[0021]　　　　Here, the data transfer unit may have a structure that includes a register which holds the size of the split region and the number n.

[0022]　　　　The data transfer unit may transfer the split region which includes all rectangular areas, from the DRAM to the memory when the access request from the data processing unit requests the rectangular areas which are adjacent or overlapping.

[0023]　　　　The data processing unit includes: a motion vector estimation unit which estimates plural motion vectors according to plural macroblocks from the inputted stream; and a decoding unit which decodes the inputted stream on a macroblock basis, and stores the decoding result in the DRAM.

A decoding sequence of the macroblocks may be changed based on the plural motion vectors so that addresses for accessing the DRAM vectors become successive.

[0024]　　　　A data access method in the present invention is a data access method for accessing a rectangular area made up of M pixels x N lines in image data from a DRAM, the DRAM having a burst mode which burst-transfers data of continuous column addresses, and storing the image data, and has an input step of inputting an access request for the rectangular area; and an address changing step of changing an access addresses included in an access request issued in the access step.　In the address converting step, addresses are converted so that a column address of data at the $(K+m)$th column, where K and m are integers and $m \leq M$, of the Lth line, and a column address of the data at the Kth column of the $(L+n)$ line, where L and n are integers and $n \leq N$, become successive.

Effects of the Invention

[0025]　　　　As above and according to the information processing device and the data access method in the present invention, there is the effect that invalid transfer loads can be reduced and the

effective bandwidth improved in comparison to data access which is shorter than the burst length of a rectangular area.

Brief Description of Drawings

[0026]   [FIG. 1]   FIG. 1 is a structural diagram of an information processing device according to the first embodiment of the present invention.

[FIG. 2]   FIG. 2 is a structural diagram of a command processing unit.

[FIG. 3]   FIG. 3 is a diagram which shows an example of a mapping of conventional address conversion.

[FIG. 4]   FIG. 4 is a diagram which shows an example of a mapping performed by an address conversion unit.

[FIG. 5]   FIG. 5 is a diagram which shows an example of the address conversion performed by the address conversion unit.

[FIG. 6]   FIG. 6 is a diagram which shows an example of the mapping performed by the address conversion unit.

[FIG. 7]   FIG. 7 (a) is a diagram which shows an example of mapping for access on a rectangular area basis, FIG. 7 (b) is a diagram which shows an example of an access sequence for a master which requests memory access on a line basis, and FIG. 7 (c) is a diagram which shows an access sequence for a master which requests memory access on a line basis.

[FIG. 8]   FIG. 8 is a structural diagram of the information processing device according to the second embodiment of the present invention.

[FIG. 9]   FIG. 9 is a diagram which shows an example of data areas transferred by a data transfer unit.

[FIG. 10]   FIG. 10 is a diagram which shows an example of a transfer area of the data transfer unit.

[FIG. 11]   FIG. 11 is a diagram which shows an example of the transfer area for the data transfer unit.

[FIG. 12]   FIG. 12 is a diagram which shows a timing chart for a decoding device.

Numerical References

[0027] 1, 3   Memory

2, 2a   Information processing device

4   Data transfer unit

10   Memory interface

11   Command processing unit

12   Data processing unit

13 to 15 Memory access request masters

20   Command bus

21   Data bus

100   Address conversion unit

201   Decoding device

202   Motion vector estimation unit

203   Decoding unit

Best Mode for Carrying Out the Invention

[0028]   Below, an information processing device according to the embodiments of the present invention is described in detail based on diagrams.

[0029]   (First embodiment)

FIG. 1 is a block diagram which shows the structure of an information processing device in the first embodiment of the present invention.   The information processing device 2 includes a memory interface 10 connected to an external memory 1, a command processing unit 11, a data processing unit 12 and masters 13, 14 and 15.   For example, in the decoding processing for a stream, the master 13 accesses the memory 1 on a rectangular area basis and the master 14 accesses the memory 1 on a line basis in order to externally output the decoded image.

[0030]   In the figure, the memory 1 is a DRAM which stores image data, having a burst mode for burst-transferring data stored in successive column addresses.

[0031]   The plural masters 13, 14 and 15 are connected to a common command bus 20 and a common data bus 21, output a command as a memory access request to the command processing unit 11, and input and output data into and from the data processing unit 12. Commands from the plural masters 13, 14 and 15 are serialized in priority rank and sent to the memory 1 through the memory interface 10. For writing, write data is sent to the memory 1 in synchronization with the sending of the command through the memory interface 10. For read data, after a command is accepted in the memory 1 and an appropriate time has elapsed, the read data is returned to data processing unit 12 through the memory interface 10. Afterwards, the read data is returned to one of the masters 13, 14 and 15 through data bus 21.

[0032]   The command processing unit 11 has an address conversion unit 100, as shown in FIG. 2, and a function for converting a logical address to a physical address. Specifically, the address conversion unit 100 converts addresses so that, in the case where a rectangular area of M pixels x N lines (M and N are integers) in the image data is accessed, the column address of data in a (K+m)th line (K and m are integers and $m \leq M$) of an Lth (L is an integer) line and the column address of data at the Kth column of a (L+n)th line (L and n are integers and $n \leq N$) become successive. FIG. 3 is an example of the mapping performed in conventional address conversion; in the case where the minimal access size of the DRAM is 4 bursts, data at one part of a (K+1)th column and all data at (K+2)th and (K+3)th columns become invalid data and the effective bandwidth drops. On the other hand, FIG. 4 is an example of mapping performed in the address conversion unit 100, when m=2 and n=1. By making the column address of data in the

(K+1)th column address and the column address of data in the (K+2)th column address successive, only a part of data in the (K+1)th column and a part of data in the (K+2)th column are wastefully transferred. In FIG. 5, an example of the logical and physical address conversion in the address conversion unit 100 is shown.

[0033] FIG. 6 is an example of the mapping of the address conversion unit 100 when m=2 and n=2.

In FIG. 6, the addresses are successive in the sequence: Lth line, (L+2)th line, (L+1)th line, (L+3)th line. When the minimum access size of the DRAM is 4 bursts, in the first access, the Lth line and (L+2)th line are accessed, and in the second access, the (L+1)th line and the (L+3)th line are accessed. For example in motion compensation processing, when the field is estimated, the access is performed by skipping a line of the reference image. Once mapping is performed as in FIG. 6, when field estimation is performed in motion compensation processing, the invalid transfer load is reduced and it becomes possible to improve the effective bandwidth. Also, when the frame is estimated, access to the lines circled 1 and circled 2 as in FIG. 6 may be repeated with no problems.

[0034] FIG. 7 is a diagram which shows the access sequence according to a master (for example the master 14), which accesses memory on a line basis. In the case where a mapping such as the one above (see FIG. 7 (a)) is performed, a master which accesses memory on a rectangular area basis and a master which accesses memory on a line basis, both of which access the same area on the DRAM, share the memory.

[0035] Although the master 14 intends to access, for example, the sequence of the circled numbers 1, 3, ..., 15, 2, 4, 6, ..., 16 (see FIG. 7 (b)) on a line basis, the master 14 may accidentally take the data of another, invalid line. In contrast, in the present

embodiment, after the command processing unit 11 and data processing unit 12 access all four lines (see FIG. 7 (c)), and once the processing of the master, which makes the line memory access request, has started, the above-mentioned inconvenience is

5   resolved.   For example, since four lines must be taken simultaneously when filter processing is performed on a decoded result and the filter process is performed using a 4-tap filter, the invalid transfer load can be reduced and the effective bandwidth can be improved when an access method such as the one above is

10   employed.

[0036]   (Second embodiment)

FIG. 8 is a block diagram which shows the structure of an information processing device in the second embodiment of the present invention.   The information processing device 2a includes a

15   memory 3 and a decoding device 201.   The decoding device 201 includes the command processing unit 11 above, data processing unit 12 above, a motion vector estimation unit 202 which estimates plural motion vectors corresponding to plural macroblocks from a stream inputted from outside, and a decoding unit 203 which

20   decodes the inputted stream on a macroblock basis and stores the decoded result in the memory 1.

[0037]   The decoding device 201 can transfer data from either the memory 1 or the memory 3.   The capacity of the memory 3 is smaller than that of the memory 1.   For example, the memory 1 is

25   a DRAM and the memory 3 is a faster SRAM.   The decoding unit 203 decodes the stream by the motion compensation process.   In the motion compensation process, the decoding unit 203 must take the decoded image data already in the memory 1 as a reference image and read out a rectangular area from the reference image.   When

30   this processing is performed, a data transfer unit 4 transfers the necessary data among the reference images stored in the memory 1 to the memory 3, not on a rectangular area basis but on a split

region basis. FIG. 9 is a diagram which shows an example of a split region transferred by data transfer unit 4. In the figure, for the sake of convenience, the area is split into 6 x 6 split regions. The example in FIG. 9 is an example which shows the reference image, which each macroblock references, at the point when the processing of 6 macroblocks has finished. Since the reference image was taken on a macroblock basis in prior art, each rectangular area on which a number is written is transferred one by one. In the present embodiment processing is performed on a basis of a number of macroblocks, and in the case where more than one reference image is referenced for a split region which has been split in advance into a size larger than the rectangular areas but smaller than the memory 3, the data transfer unit 4 transfers the split region. For example, split regions shown in FIG. 9 such as 1000, 1001 and 1002 are transferred. For 1002, when the area split in advance is less than half of the capacity of the memory 3, two rectangles may be transferred simultaneously. Also, a data transfer sequence for 1000, 1001 and 1002 is not of concern. Note that the split size, which is decided in advance, may be made alterable in a register. This register may be changed dynamically. Since it becomes possible to take a large rectangular area once, the invalid transfer area may be reduced and it becomes possible to improve the effective bandwidth.

[0038]   FIG. 10 is a diagram which shows an example of another transfer area in data transfer unit 4. The data transfer unit 4 performs data transmission when the predetermined number of areas are referred to within a range below the capacity of the memory 3. Note that the information processing device may be set so that the number of areas which are decided in advance may be alterable by the register setting within data transfer unit 4. The register may be changed dynamically. Since the split region, which is larger than the rectangle areas in one transfer, is transferred, the

invalid transfer area is reduced by multiple transfers and it becomes possible to improve the effective bandwidth.

[0039]   FIG. 11 is a diagram which shows still another transfer area example for data transfer unit 4.   Data transfer unit 4 transfers plural reference images simultaneously, only when the reference image addresses are successive or where the reference images overlap in the range below the capacity of the memory 3. Since it becomes possible to take a large rectangular area once, the invalid transfer area can be reduced and it becomes possible to improve the effective bandwidth.

[0040]   FIG. 12 is a diagram which shows a timing chart of the decoding device 201.   As pictured, there are roughly four processes in the decoding process.   Among these processes, the motion vector estimation unit 202 analyzes the input stream as a PreSearch process and performs a process in advance to estimate plural motion vectors in the plural macroblocks.   In the PreSearch process, after estimating the motion vectors in plural macroblocks, the decoding unit 203 changes the decoding order for the macroblock so that the addresses of the rectangle areas which must be referenced become successive, based on the value of the motion vectors estimated; next, the VLD (Variable-length Decoding), the IDCT (Inverse Discreet Cosine Transformation) and the MC (Motion Compensation) processes are performed in pipelines.   In this case, the motion compensation process takes reference images out of the memory in the sequence in which the macroblocks arrive, and since the order of the macroblocks is transposed in advance so that the addresses of the reference images become successive, and plural rectangular areas come to be included in one transfer of the split regions, the invalid transfer area can be reduced and it becomes possible to improve the effective bandwidth.

Industrial Applicability

[0041]   The present invention is applicable to an information processing device which includes a DRAM that stores image data and has a burst mode, which burst-transfers successive data at successive column addresses, and applies to an information processing device, for example a network terminal which reproduces a compression-encoded stream, a digital camera, a cellular telephone with camera, a DVD recorder/player, a PDA, a personal computer and so on which are capable of recording and playing back DTV (digital television) movies.